DOCUMENT RESUME

ED 194 527 TM 800 215

AUTHOR Powers, Stephen: Gallas, Edwin J.
TITLE The Factorial Structure and Item Composition of the Comprehensive Tests of Basic Skills, Reading Comprehension Test, When Administered to ESEA Title I Sixth and Seventh Grade Students.
INSTITUTION Tucson Unified School District, Ariz.
PUB DATE Apr 80
NOTE 26p.: Paper presented at the Annual Meeting of the American Educational Research Association (64th, Boston, MA, April 7-11, 1980).

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Achievement Tests: Compensatory Education: Factor Analysis: *Factor Structure: Grade 6: Grade 7: Intermediate Grades: *Item Analysis: Junior High Schools: *Reading Comprehension: Reading Difficulties: *Reading Tests: *Test Validity
IDENTIFIERS *Comprehensive Tests of Basic Skills: *Elementary Secondary Education Act Title I

ABSTRACT

The general purposes of this study were to investigate the factorial validity of the Comprehensive Tests of Basic Skills (CTBS), Reading Comprehension subtest, and to study the item difficulties in order to determine areas of strengths and weaknesses of sixth and seventh grade Elementary Secondary Education Act (ESEA) Title I students. A sample of 581 Title I students was administered the test. Using Bartlett's test, it was determined that the item intercorrelation matrices were worth factor analyzing. The CTBS did not have a strong factor structure. Rather than being unidimensional, the test results appeared to be multidimensional. This study suggests that the results from testing Title I students with standardized tests should be interpreted cautiously. The total score for each student is far from representing any one construct and the error variance in the testing is considerable. (Author/BW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED194527

U S DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

THE FACTORIAL STRUCTURE AND ITEM COMPOSITION OF THE COMPREHENSIVE TESTS OF BASIC SKILLS, READING COMPREHENSION TEST, WHEN ADMINISTERED TO ESEA TITLE I SIXTH AND SEVENTH GRADE STUDENTS

by

Stephen Powers, PhD
Title I Research Specialist

Data Analysis--Edwin J. Gallas

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Powers, Stephen
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Department of Research and Evaluation
Director, Barbara S. Prentice, PhD
Tucson Unified School District
Tucson, Arizona

Paper presented at the Annual Meeting of the American Educational Research Association, Boston, Massachusetts, April 7-11, 1980.

TM800 215

The Factorial Structure and Item Composition of the Comprehensive Tests of Basic Skills, Reading Comprehension Test, When Administered to ESEA Title I Sixth and Seventh Grade Students

Porter, Schmidt, Floden and Freeman (1978) advocate an "intimate familiarity" with the measures used to estimate a program's effect. This familiarity would assist in the match of test content with program objectives. The need for program valid achievement testing has been emphasized by many authors (Porter, et al, 1978; Bloom, Hastings and Madaus, 1971; Nunnally and Wilson, 1975; and Shoemaker, 1975).

One approach to obtaining an intimate familiarity with an achievement test, is to scrutinize the items and investigate the factorial validity of the test. Achievement tests are purportedly constructed to be unidimensional and this is supported by high internal reliability estimates. One empirical definition of a test's unidimensionality is a large first factor on the item intercorrelation matrix (Porter, et al, 1978). A careful examination of achievement tests are especially crucial when such tests are administered to low-achieving Title I students.

The general purposes of the present study are to investigate the factorial validity of the Comprehensive Tests of Basic Skills (CTBS), Reading Comprehension Test, Level 2, Form S, and to study the item difficulties in order to determine areas of strengths and weaknesses of sixth and seventh grade Title I students. Specifically, the focus of this study is on the following research questions involving Title I sixth and seventh grade students:

1. Is the intercorrelation matrix which is obtained from the items of the CTBS Reading Comprehension Test worth factor analyzing?

2. Does the CTBS Reading Comprehension Test have a strong factor structure?

3. Is the CTBS Reading Comprehension Test unidimensional?

4. Does one factor account for most of the intercorrelations of the items of the CTBS Reading Comprehension Test?

5. What are the variance components of the CTBS Reading Comprehension Test?

6. What are the item difficulties of the CTBS Reading Comprehension Test when administered to Title I students?

7. Are there certain areas where Title I students appear to be especially weak when compared with CTBS norms?

Method

In a large southwestern school district, 581 ESEA Title I students (220 sixth grade students and 361 seventh grade students) were administered the CTBS, Reading Comprehension Test, Level 2, Form S, in the spring of 1979. All students were enrolled in a Title I program in which students were selected principally because of poor reading performance in stanines 1-3 on a norm-referenced test. The ethnic background of participants is given in Table 1.

Table 1. Ethnic Backgrounds of Participants

Ethnic Background	Percent
American Indian	6
Black	10
Hispanic	75
Other	9
	100

The CTBS/S, Levels 1-4, is a battery of seven tests measuring three basic skills areas: Reading, Language and Mathematics (CTBS/S Test Coordinator's Handbook 1976). The skills areas were classified using Bloom's Taxonomy of Educational Objectives. In the test

development, efforts were made to reduce racial and ethnic bias. The K-R 20 reliabilities at each grade level for vocabulary, comprehension and total scores are almost all above .90 with standard errors of measurement from .25 to 1.01 in grade equivalent units. Moreover, it appeared that systematic procedures were followed in test development to ensure content validity. The CTBS/S Reading Comprehension Tests, Levels 1-4 are composed of 45 items and each item in Levels 1-4 contains a multiple choice involving four alternatives.

Before the statistical analysis, raw scores were converted to the CTBS/S Expanded Standard Score Scale. Calculations for the present study were performed with expanded standard scores unless otherwise noted. As the expanded standard scores form a normalized scale with assumed equal intervals, it was believed this metric was more appropriate for statistical analysis. This choice of a metric is in agreement with the technical advice to ESEA Title I evaluators (Tallmadge and Wood 1976). When it was desired to convert an out-of-level test statistic up to the appropriate on-level statistic, the raw scores were converted to expanded standard scores and the desired statistic was computed with expanded standard scores. Then, the appropriate grade level table was referred to, and by means of the appropriate table, the expanded standard score was used to ascertain the grade level, raw score, percentile, stanine or grade equivalent.

Results

Participants in this study were selected for Title I because of poor reading ability. The guessing level is often high with Title I students. Moreover, research has shown that often random data can produce interpretable results when it is factor analyzed (Armstrong

and Soelberg, 1968; Horn, 1967; Humphreys, Ilgen, McGrath, and Montanelli, 1969). Therefore, the first question to be answered in this study was whether the item intercorrelation matrix was worth factor analyzing. Gorsuch (1976) presents Bartlett's (1950) test of significance for a correlation matrix:

$$\chi^2 = -\left(n - 1 - \frac{2v + 5}{6}\right) \text{Log}_e |R|$$

where
- n = number of individuals
- v = number of variables
- $|R|$ = the determinant of the correlation matrix

with $df = v\,\frac{v - 1}{2}$

The results for the sixth grade are $\chi^2 = 1483$, df = 990 and for the seventh grade $\chi^2 = 2253$, df = 990. These large chi-square values can be converted to a standard normal variable using this formula appearing in Edwards (1968, p. 67), and Glass and Stanley (1970, p. 520):

$$Z = \sqrt{2\,(\chi^2)} - \sqrt{2(df) - 1}$$

The chi-square values converted to standard normal variables are: for the sixth grade Z = 9.98 and for the seventh grade Z = 22.64. Since these two values were greater than Z = 1.96, $\alpha = .05$, the two intercorrelation matrices could legitimately be factored.

To investigate the factor structure of the CTBS Reading Comprehension Test when administered to Title I sixth and seventh grade students a principal components factor analysis was conducted followed by a Varimax rotation. Principal components with eigenvalues less than 1.00 were deleted from further study. Factor loadings greater than or equal to .30 were identified for further study. Since this was an

exploratory factor analysis with more than twenty variables, the recommendation of Nunnally (1978, p. 418) was followed; ". . . it is strongly recommended that unities be placed in the diagonal spaces and that component analysis be undertaken." Nunnally further recommends that the principal components analysis be followed by Varimax rotation (p. 418). The results of the components analysis are presented in Tables 2-5.

Table 2. Principal-Components Factor Analysis of the Correlation Matrix of the Comprehensive Tests of Basic Skills, Reading Comprehension Test Administered to ESEA Title I Sixth Grade Students*

Factor	Eigenvalue	Percent of Variance	Cumulative Percent of Variance
1	4.42	9.8	9.8
2	2.08	4.6	14.4
3	1.85	4.1	18.5
4	1.67	3.7	22.3
5	1.59	3.5	25.8
6	1.57	3.5	29.3
7	1.54	3.4	32.7
8	1.45	3.2	35.9
9	1.38	3.1	39.0
10	1.34	3.0	42.0
11	1.25	2.8	44.7
12	1.21	2.7	47.4
13	1.18	2.6	50.1
14	1.14	2.5	52.6
15	1.10	2.4	55.1
16	1.09	2.4	57.5
17	1.04	2.3	59.8
18	1.01	2.2	62.0
19	1.01	2.2	64.3

*N = 220

Table 3. Factor Loadings* of the Principal Components Factor Analysis of the Correlation Matrix of the Comprehensive Tests of Basic Skills, Reading Comprehension Test, Administered to ESEA Title I Sixth Grade Students (N = 220)

Factor	Item	Loading
Factor 1	7	68**
	21	46
	22	43
	24	38
	37	-44
Factor 2	2	68
	5	72
Factor 3	41	80
	42	40
Factor 4	12	33
	19	45
	21	35
	23	59
	26	54
	28	35
	29	72
Factor 5	1	64
	6	31
	12	31
	16	33
	36	63
Factor 6	14	57
	32	76
Factor 7	18	76
	24	36
Factor 8	12	-34
	17	75
	30	33
	38	45
	42	32
Factor 9	6	49
	8	66
	11	61
	16	35
	45	32
Factor 10	22	-32
	40	47
	43	76
Factor 11	4	72
	10	61
	28	-36
Factor 12	27	66
	30	42
	31	68
Factor 13	9	82
	10	40
	24	-46
Factor 14	35	77
	36	33
	45	35
Factor 15	15	79
	39	-37
	42	30
Factor 16	3	74
	13	50
Factor 17	28	48
	33	80
Factor 18	19	32
	39	49
	44	69
Factor 19	25	87

* Only Factor loadings ≥ 30
** Decimal points are omitted.

Table 4. Principal Components Factor Analysis of the Correlation Matrix of the Comprehensive Tests of Basic Skills, Reading Comprehension Test, Administered to ESEA Title I Seventh Grade Students*

Factor	Eigenvalue	Percent of Variance	Cumulative Percent of Variance
1	5.74	12.8	12.8
2	2.25	5.0	17.8
3	1.74	3.9	21.6
4	1.45	3.2	24.8
5	1.39	3.1	27.9
6	1.34	3.0	30.9
7	1.32	2.9	33.8
8	1.26	2.8	36.6
9	1.22	2.7	39.3
10	1.20	2.7	42.0
11	1.17	2.6	44.6
12	1.13	2.5	47.1
13	1.10	2.4	49.6
14	1.07	2.4	51.9
15	1.04	2.3	54.2
16	1.02	2.3	56.5
17	1.00	2.2	58.7

*N = 361

Table 5. Factor Loadings* of the Principal Components Factor Analysis of the Correlation Matrix of the Comprehensive Tests of Basic Skills, Reading Comprehension Test, Administered to ESEA Title I Seventh Grade Students (N = 361)

Factor	Item	Loading
Factor 1	3	32**
	5	56
	6	52
	7	35
	8	55
	11	31
	14	40
	20	63
	21	48
	22	31
	26	32
Factor 2	40	73
	41	41
	42	71
Factor 3	2	-34
	7	32
	9	54
	10	68
	38	41
Factor 4	3	32
	18	37
	37	81
	45	39
Factor 5	26	47
	28	55
	30	64
	31	35
	34	55
	38	41
Factor 6	3	-30
	43	68
	44	38
	45	36
Factor 7	3	37
	4	32
	6	31
	13	67
	17	51
Factor 8	1	81
	2	35
Factor 9	19	55
	27	64
	39	46
Factor 10	7	31
	33	71
	36	51
Factor 11	11	40
	18	33
	22	60
	23	73
Factor 12	25	78
	26	30
	45	41
Factor 13	14	55
	15	44
	19	31
	29	63
	31	57
Factor 14	7	34
	9	32
	35	72
Factor 15	9	33
	17	32
	24	68
Factor 16	12	77
	41	38
Factor 17	16	47
	32	79

* Only Factor Loadings $\geq$ 30
** Decimal points are omitted.

A strong factor structure was explained by Nunnally (1978, p. 418):

> By a strong structure is meant one in which there are obvious groupings of the variables, as evidenced in substantial correlations among members of the groups and much lower correlations between members of different groups. In terms of factor analysis products, this means that each factor has some variables that load mainly on it alone, and there are at least four tests with loadings above .50. Unless a factor is at least that strong, it would be best to ignore it.

From an inspection of Table 3, in which the factor loadings for the sixth grade students are displayed, only factors 1 and 4 approach the criterion for strong factor structure. In Table 5 in which the results for seventh grade students are presented, factor 1 clearly meets the criterion of strong factor structure and factors 5 and 13 approach the criterion. Thus, from the principal components analysis of the correlation matrices of the sixth and seventh grade Title I students, only factor 1 of the seventh grade students clearly attains the criterion for strong factor structure.

Henrysson (1971, pp. 153-4) spoke of the use of factor analysis to determine the dimensionality of a test: A factor analysis

> is a check of the dimensionality of the test. If only one large general factor is found, except for some other very small loadings, the test is measuring only one main dimension.

The principal components factor analysis extracted nineteen factors from the sixth grade intercorrelation matrix and seventeen factors from the seventh grade intercorrelation matrix with the restriction that eigenvalues be greater than or equal to 1.00. These results suggest that the correlation matrix of the CTBS Reading

Comprehension Test results can be partitioned into many uncorrelated dimensions when administered to sixth and seventh grade Title I students. With Title I students the CTBS Reading Comprehension Test did not appear to be unidimensional, rather it appeared to be factorially complex.

For construct validity, a test might be expected to have one factor account for most of the variance of the items. According to Cronbach (1971),

> For construct validity, one would ordinarily desire sufficient coherence among items contributing to a score so that one factor would account for most of the inter-correlation of items. But many constructs are exceptions. If the construct is conceived as a conglomerate, one would expect the items to fall into clusters. (p. 470)
>
> Tests that by hypothesis are indicators of a certain construct are expected to show substantial loadings on the same factors. When one of them loads on a second factor, this shows that the indicator is impure. (p. 469)

The first factor extracted from the correlation matrix of the sixth grade students accounted for 9.8 percent of the total variance of the test. The remaining 54.5 percent of the variance of the test was distributed over the remaining 18 factors. The first factor extracted from the correlation matrix of the seventh grade students accounted for 12.8 percent of the total test variance. The remaining 48.9 percent of the variance was accounted for by the remaining sixteen factors. For both sixth and seventh grade students, the partitioning of variance among so many factors indicates a lack of coherence among items. Moreover, it appeared that one factor would not account for most of the intercorrelation of items.

Kerlinger (1973, p. 470) presents the three variance components of measurement theory: Total variance (V_t), a common factor variance (V_{co}), specific variance (V_{sp}) and error variance (V_e). Common factor variance is that variance that two or more measures share in common, and specific variance is that variance not shared by any other measure. The relationship between these components of total variance are specified in the following equation:

$$V_t = V_{co} + V_{sp} + V_e$$

This theory may be applied to the present study where the variances are proportions of the total variance. For the CTBS Reading Comprehension Test, the KR Formula 20 for sixth grade results was .75 and for seventh grade was .82. The total common factor variance extracted from the sixth grade data was 64.3 percent and the seventh grade data was 58.7 percent. Since the reliability is equal to the common factor variance plus the specific variance, the basic equation partitioning the total variance into common factor variance, specific variance, and error variance can be written:

Sixth Grade Data

100 % = 64.3% + 10.7% + 25%

Seventh Grade Data

100 % = 58.7% + 23.3% + 18%

In the Comprehensive Tests of Basic Skills, items are classified according to process (Bloom's Taxonomy) and content. Table 6 presents the item classification for the CTBS, Reading Comprehension Test,

Level 2, Form S. Description of the process categories are provided in Appendix A, and a description of the content is in Appendix B.

The item difficulties (p-values) for the sixth and seventh grade were subtracted from the CTBS p-value norms. The averages of the differences were computed to get an overall estimate of the sixth and seventh grade differences with the norms. Within each of the seven process-content categories, the averages of the differences were computed. Tables 7 and 8 presents those differences. The higher the number, the greater the discrepancy with the national norms of that group of items.

Table 6. Item Classification for Comprehensive Tests of Basic Skills, Reading Comprehension Test, Level 2, Form S

Process-Content Category	Items
1. Literal Recall	6, 13, 16, 21, 22, 26, 33, 34
2. Rewording	8, 20, 23, 26, 41, 45
3. Context Clues	7, 9, 14, 15, 28
4. Main Idea	2, 12, 24, 38, 44
5. Descriptive Words	19, 30, 32, 39
6. Conclusions	1, 10, 11, 17, 31, 37, 42, 43
7. Structure/Style	3, 4, 5, 18, 25, 27, 29, 35, 40

Table 7. Item Difficulty Values of Students in Grades 6 and 7 Compared with CTBS National Norms

Items	Grade 6	Grade 7	Norms	N-6	N-7	Average of Differences
1	55	64	76	21	12	15.4
2	55	68	80	25	12	16.9
3	62	72	81	9	19	15.2
4	46	40	66	20	26	23.7
5	56	71	82	26	11	16.7
6	63	85	87	24	2	10.3
7	39	50	77	39	27	31.5
8	46	66	82	36	16	23.6
9	30	38	69	39	31	34.0
10	33	43	69	36	26	29.8
11	56	64	81	25	17	20.0
12	32	43	43	11	--	4.2
13	56	62	77	21	15	17.3
14	42	57	69	27	12	17.7
15	40	57	74	34	17	23.4
16	58	68	71	13	3	6.8
17	43	50	58	15	8	10.7
18	23	36	55	32	19	23.9
19	50	56	69	19	13	15.3
20	50	70	79	29	9	16.6
21	58	73	83	25	10	15.7
22	46	69	77	31	8	16.7
23	52	65	74	22	9	13.9
24	46	47	65	19	18	18.4
25	29	27	52	23	25	24.2
26	64	80	86	22	6	12.1
27	25	33	56	31	23	26.0
28	59	69	72	13	3	6.8
29	26	42	62	36	20	26.1
30	53	64	75	22	11	15.2
31	33	37	64	31	27	28.5
32	14	19	39	25	20	21.9
33	33	41	64	31	23	26.0
34	42	48	65	23	17	19.3
35	30	46	61	31	14	20.4
36	27	36	59	32	23	26.4
37	19	22	32	13	10	11.1
38	26	39	66	40	27	31.9
39	21	34	58	37	24	28.9
40	30	41	51	21	10	14.2
41	35	42	65	30	23	25.7
42	35	43	56	21	13	16.0
43	29	35	54	25	19	21.3
44	26	37	58	32	21	25.2
45	22	24	48	26	24	24.8

Item difficulties (p-values) have been converted to percentages.

Table 8. The Average of the Mean Differences Computed Within Each CTBS Content Category

Content Category	Mean Differences
1. Literal Recall	.15
2. Rewording	.22
3. Context Clues	.23
4. Main Idea	.19
5. Descriptive Words	.20
6. Conclusions	.19
7. Structure/Style	.21

In Table 7, the average of the p-values differences are computed to identify categories which are easier or more difficult for Title I sixth and seventh grade students. This data is presented as a bar graph in Figure 1. The Literal Recall Category was the easiest when compared with national norms. The average p-value difference was .15. This category is the recall of facts and details explicitly stated in a passage.

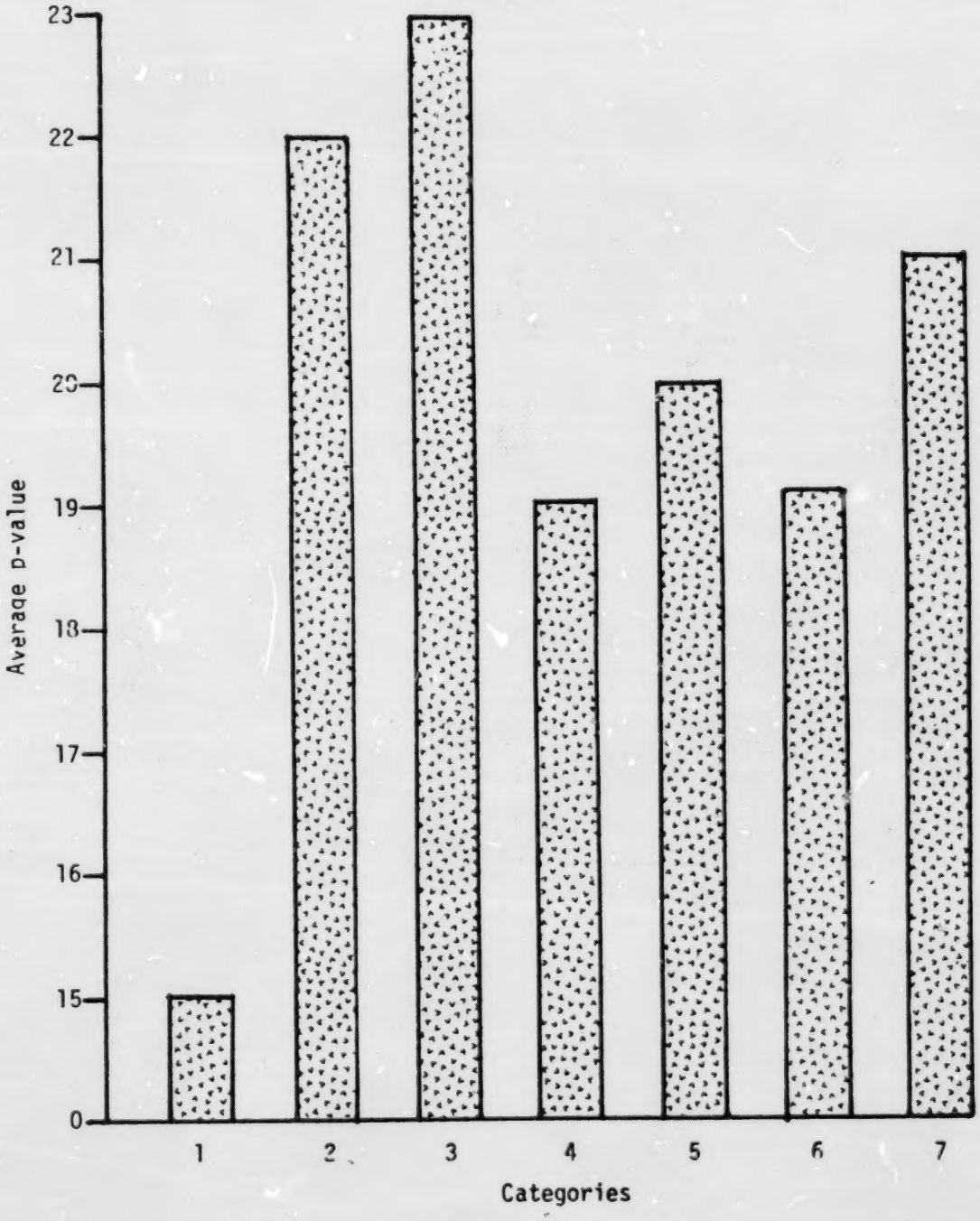

Figure 1. Average mean differences of combined sixth and seventh grade students.

The categories of moderate difficulty were Main Idea (.19), Descriptive Words (.20) and Conclusions (.19). Main Idea refers to such skills as identifying the topic of a passage, restating or summarizing the central thought of a passage and identifying the authors purpose. Descriptive Words include defining the physical attributes of an objective or person presented in a passage and analyze emotions experienced in the passage. Conclusions include skills such as drawing conclusions based on facts stated or implied in the passage and perceived cause and effect relationships.

The difficult categories when compared to national norms were Rewording (.22), Context Clues (.23) and Structure/Style (.21). Rewording is to give facts and details in a passage and to answer questions by choosing paraphrased or reworded material. Context Clues refers to determining word meaning from the context in which a word is used. Structure/Style includes the identification of the writers use of words, figurative language and the interpretation of symbols. Moreover, this category refers to understanding structured devices that help convey meaning such as punctuation, function words or tense.

The CTBS content-categories were not ascending in the average difficulties within categories. Rather one category, Literal Recall was noticeably easier for combined sixth and seventh grade Title I students. The other six categories were classified as moderately difficult to difficult for Title I students.

A z-statistic was computed to test the significance of the difference between two independent proportions (Glass and Stanley, 1970, pp. 324-326). This statistical test was undertaken to identify those p-values that were significantly different between the sixth and seventh grade students. The results are presented in Table 9.

Table 9. Comparison of Item Difficulty Values of the Sixth and Seventh Grade Title I Students

Items	Grade 6	Grade 7	z-value
1	.55	.64	-2.09*
2	.55	.68	-3.06**
3	.62	.72	-2.44**
4	.46	.40	1.38
5	.56	.71	-3.58**
6	.63	.85	-5.94**
7	.38	.50	-2.73**
8	.46	.66	-4.61**
9	.30	.38	-1.90*
10	.33	.43	-2.32*
11	.56	.64	-1.86*
12	.32	.43	-2.56**
13	.56	.62	-1.39
14	.42	.57	-3.41**
15	.40	.57	-3.86**
16	.58	.68	-2.37**
17	.43	.50	-1.59
18	.23	.36	-3.18**
19	.50	.56	-1.37
20	.50	.70	-4.70**
21	.58	.73	-3.64**
22	.46	.69	-5.34**
23	.52	.65	-3.02**
24	.46	.47	- .23
25	.29	.27	.51
26	.64	.80	-4.15**
27	.25	.33	-1.98*
28	.59	.69	-2.39**
29	.26	.42	-3.77**
30	.53	.64	-2.55**
31	.33	.37	- .95
32	.14	.19	-1.50
33	.33	.41	-1.87*
34	.42	.48	-1.37
35	.30	.46	-3.70**
36	.27	.36	-2.17*
37	.19	.22	- .84
38	.26	.39	-3.11**
39	.21	.34	-3.24**
40	.30	.41	-2.58**
41	.35	.42	-1.62
42	.35	.43	-1.85*
43	.29	.35	-1.45
44	.26	.37	-2.65**
45	.22	.24	- .54

* $p < .05$
** $p < .01$

Using a one-tailed test of significance twenty-four items differed significantly between sixth and seventh grade students ($p < .01$), eight items difference significantly between sixth and seventh grade students ($p < .05$) and thirteen items did not differ significantly. Items that were different between sixth and seventh grade Title I students did not fall into easily interpretable categories.

Discussion

Using Bartlett's test, it was determined that the Title I sixth and seventh grade item intercorrelation matrices were worth factor analyzing. This was a necessary preliminary to this factor analytic investigation of test results of low-achieving students. Clearly the CTBS Reading Comprehension Test did not have a strong factor structure when administered to Title I students. Rather than being unidimensional, the test results appeared to be multidimensional. The test results were factorially complex. One factor did not account for most of the intercorrelations among the items. Large specific and error components were found in the test results. This study presents evidence that when the CTBS Reading Comprehension Test is administered to low-achieving Title I students, the results are not measuring a single, unidimensional construct.

The error variance is considerable in these test results. Since principal components analysis totally partitions the test variance, not just the common factor variance, with the only restriction that eigenvalues be greater than or equal to 1.00, the error and specific variance was probably intermixed with factors and this accounted for the large number of factors.

This study suggests that the results from testing Title I students with standardized tests should be interpreted cautiously. The total score for each student is far from representing any one construct and the error variance in the testing is considerable. One method suggested for reducing error variance is out-of-level testing. It should be pointed out that the seventh grade students were tested out-of-level, that is, one level below that recommended by the publisher, and the sixth grade students were tested on level. The factor analytic results were not noticeably different for these two groups. One might consider replicating this method with more extreme groups.

In summary, two methods of scrutinizing test results is with factor analysis and item analysis. Researchers should not rely simply on interpretations of total scores, means, and mean differences. In order to more fully understand the mechanisms of the test, one should explore the test results in a variety of ways. Two of those ways are factor analysis and item analysis.

REFERENCES

Armstrong, J. S., and Soelberg, P. On the Interpretation of Factor Analysis, Psychological Bulletin, 70(5), p. 361, 1968.

Bartlett, M. S. Tests of Significance in Factor Analysis. British Journal of Psychology, 3(2), p. 77, 1950.

Bloom, B. S., Hastings, T. M., and Madaus, G. F. Handbook on Formative and Summative Evaluation of Student Learning. New York: McGraw-Hill, 1971.

Cronbach, L. J. Test Validation. In R. L. Thorndike (Ed.), Educational Measurement (2nd ed.), Washington, D.C.: American Council on Education, 1971.

Gorsuch, R. L. Factor Analysis Philadelphia: Saunders, 1974.

Harman, H. H. Modern Factor Analysis (3rd ed.). Chicago: The University of Chicago Press, 1976.

Henrysson, S. Gathering, Analyzing, and Using Data on Test Items. In R. L. Thorndike (Ed.), Educational Measurement (2nd ed.), Washington, D.C.: American Council on Education, 1971.

Humphreys, L. G., Ilgen, D., McGrath, D., and Montanelli, R. Capitalization on Change in Rotation of Factors. Educational and Psychological Measurement, 29(2), p. 259, 1969.

Kerlinger, F. N. Foundations of Behavioral Research. (2nd ed.), New York: Holt, Rinehart and Winston, 1973.

Nunnally, J. C. Psychometric Theory (2nd ed.). New York: McGraw-Hill, 1978.

Nunnally, J. C., and Wilson, W. H. Method and Theory for Developing Measures in Evaluation Research. In Struening and Guttentag (Ed.), Handbook of Evaluation Research, Beverly Hills: Sage, 1975.

Porter, A. C., Schmidt, W. H., Floden, R. E., and Freeman, D. J. Practical Significance in Program Evaluation. American Educational Research Journal, Fall 1978, 15(4), 529-539.

Shoemaker, D. M. Toward a Framework for Achievement Testing. Review of Educational Research, 1975, 45, pp. 127-147.

Thorndike, R. L. Correlational Procedures for Research. New York: Gardner Press, 1978.

APPENDIX A

Process Categories Adapted from Bloom's Taxonomy Used for Classifying Items in the Comprehensive Tests of Basic Skills, Reading Comprehension Test

Recognition

Recognize, or recall and apply, facts, theories, laws, concepts, structures, methods, and formulas previously learned.

Translation

Transform concepts, terms, and symbols into equivalent forms, as in interpreting graphically presented data.

Interpretation

Understand facts and concepts in, and make inferences from, written material. Associate data and ideas and comprehend relationships stated or implied in material. Summarize and restate significant themes or concepts.

Application

Apply prior knowledge of facts or techniques to answer questions and solve problems. Utilize previously learned research skills and techniques to determine function or meaning of new material.

Analysis

Apply formal logic to the breakdown of material into organizational patterns or principal components; the determination of validity, bias, tone; the hierarchical arrangement of facts; the cognizance of total problems.

APPENDIX B

Description of Content Categories for the Comprehensive Tests of Basic Skills, Reading Comprehension Test

Literal Recall

Recall facts and details explicitly stated in a passage. The student is required to answer "who," "what," "where," and "when" questions.

Rewording

Given facts and details in a passage, answer questions by choosing paraphrased or reworded material.

Context Clues

Determine word meaning from the context in which a word is used.

Main Idea

Identify the topic of a passage, restate or summarize the central thought in a passage, identify the author's purpose in writing the passage, derive a lesson or moral from the passage, and select the best title for a passage.

Descriptive Words

Define the physical attributes of an object or person presented in the passage, analyze emotions experienced in the passage, or decide what quality best characterizes a person.

Conclusions

Draw conclusions based on facts stated or implied in the passage or perceive cause-and-effect relationships between events or ideas.

Structure/Style

Identify the writer's use of words, including figurative language, to evoke a feeling or create an image, interpret symbols, and understand the writer's point of view and method of conveying meaning, such as punctuation, function words, or tense.